Intelligent methods to identify incorrect reviews in cloud reputation systems

Quynh Ngoc Thuy Do

Supervisor: Associate Professor Farookh Khadeer Hussain

Co-supervisor: Doctor Christy Jie Liang

University of Technology Sydney

Faculty of Engineering and Information Technology

Centre for Artificial Intelligence

25th August, 2019

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Quynh Ngoc Thuy Do declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy degree, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: 25th August 2019

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my principal supervisor, Associate Professor Farookh Khadeer Hussain for his support, kindness, gratitude, motivation and wisdom. Without his help, I could not have completed this journey. He not only guided my research and provided advice on writing the thesis and academic papers, he also provided guidance in every aspect of life. Words are not enough to thank him for everything he has done for me. Secondly, I would like to express my appreciation to my co-supervisor, Doctor Christy Jie Liang, who supported me until the very end of my PhD journey.

I would like to thank Dong Nai province, Vietnam for awarding me a scholarship. Many thanks to the Centre for Artificial Intelligence, Faculty of Engineering and Information Technology for awarding me travel grants to present my research work.

I would like to thank my friends and colleagues: Khimmy, Raza, Asma, Walayat, Ikram and Yahya with whom I had a memorable time at UTS. I would like to thank Ms Michele Mooney for her services in proofreading this thesis.

Finally, I would like to thank my family, without whom I would not be the person I am today. To my wonderful parents, Mr Nghia Do and Mrs Loan Dinh, who have always sacrificed a lot for me and who have always been there to support me. I would like to thank my husband, Bang Nguyen and my two little children, Quynh-An Nguyen and The-An Nguyen for their unwavering support and sacrifices made for me. Thanks also to my uncle Tien Dinh, my brother Thac Do, my sister-in-law Lan Nguyen and my nephew Viet Do for their kind words and unwavering belief in me.

Quynh Ngoc Thuy Do

LIST OF PAPERS/ PUBLICATIONS INCLUDED

Do, Q.N.T., Zhilin, A., Junior, C.Z.P., Wang, G. & Hussain, F.K., 2016, 'A network-based approach to detect spammer groups', *2016 International Joint Conference on Neural Networks (IJCNN),* IEEE, pp. 3642-8.

Do, Q.N.T., Hussain, F.K. & Nguyen, B.T., 2017, 'A fuzzy approach to detect spammer groups', *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, pp. 1-6.

Table of Contents

CERTIFICATE OF ORIGINAL AUTHORSHIP i

ACKNOWLEDGEMENT ii

LIST OF PAPERS/ PUBLICATIONS INCLUDED iii

LIST OF FIGURES ix

LIST OF TABLES xi

ABSTRACT xii

THESIS SUMMARY xiii

CHAPTER 1: INTRODUCTION 1

1.1. Introduction 1

1.2. The importance of identifying incorrect reviews in cloud reputation systems 1

1.3. Issues related to identifying incorrect reviews 3

1.4. Objectives of the thesis 5

1.5. Scope of the thesis 6

1.6. Significance of the thesis 6

1.7. Plan of the thesis 7

1.8. Conclusions 8

1.9. References 9

CHAPTER 2: LITERATURE REVIEW 11

2.1. Introduction 11

2.2. Definitions of incorrect reviews in the literature 11

2.2.1. Definition of ballot stuffing 11

2.2.2. Definition of bad mouthing 12

2.2.3. Definition of spammer groups 12

2.2.4. Definition of cliques 13

2.3. Approaches to identify ballot stuffing and bad mouthing 14

2.3.1. Internal factor-based methods to identify ballot stuffing and bad mouthing 16

2.3.2. External-based methods to identify ballot stuffing and bad mouthing 17

2.4. Approaches to identify spammer groups 20

2.4.1. Linguistic-based methods to identify spammer groups....20
2.4.2. Behaviour-based methods to identify spammer groups....21
2.4.3. Network-based methods to identify spammer groups....22
2.5. Approaches to identify cliques....24
2.6. Critical evaluation of existing approaches: an integrative view....26
2.6.1. Ballot stuffing and bad mouthing....26
2.6.2. Spammer groups....27
2.6.3. Cliques....27
2.6.4. Shortcomings of the existing literature on identifying incorrect reviews....27
2.7. Conclusions....29
2.8. References....30
CHAPTER 3: PROBLEM DEFINITION....33
3.1. Introduction....33
3.2. Key terms and concepts....33
3.2.1. Cloud computing....33
3.2.2. Reputation system....34
3.2.3. Incorrect review....34
3.2.4. Ballot stuffing....34
3.2.5. Bad mouthing....35
3.2.6. Spammer group....35
3.2.7. Clique....35
3.2.8. Internal factor-based methods....35
3.2.9. External factor-based methods....35
3.2.10. Linguistic-based method....35
3.2.11. Behaviour-based method....35
3.2.12. Network-based method....36
3.2.13. Agent....36
3.3. Problem definition....36
3.4. Research questions....38
3.5. Research objectives....38
3.5.1. Research Objective 1: Propose an intelligent method for identifying ballot stuffing in cloud reputation systems.....39

3.5.2. Research objective 2: Propose an intelligent method to identify bad mouthing in cloud reputation systems. 40
3.5.3. Research objective 3: Propose an intelligent method to identify spammer groups in cloud reputation systems 41
3.5.4. Research objective 4: Define the concept of cliques and propose an intelligent method to identify cliques in cloud reputation systems 42
3.5.5. Research issue 5: Validate the proposed methods 42
3.6. The research approach to problem solving 43
3.7. Conclusion 44
3.8. References 44
CHAPTER 4: SOLUTION OVERVIEW 46
4.1. Introduction 46
4.2. Overview of the proposed solutions 46
4.3. Overview of the solution for identifying ballot stuffing 47
4.4. Overview of the solution for identifying bad mouthing 49
4.5. Overview of the solution for identifying spammer groups 51
4.6. Overview of the solution for identifying cliques 53
4.7. Overview of the experiment 55
4.8. Conclusion 56
4.9. References 56
CHAPTER 5: NETWORK-BASED METHOD TO IDENTIFY BALLOT STUFFING 57
5.1. Introduction 57
5.2. Solution to identify ballot stuffing 58
5.2.1. Step 1: Calculate the suspicion score of reviewers 58
5.2.2. Step 2: Calculate the suspicion score of a review 59
5.2.3. Step 3: Calculate the suspicion score of product k 60
5.2.4. Step 4: Apply naïve Bayes to identify if the review is ballot stuffing or not 62
5.3. Conclusion 63
CHAPTER 6: NETWORK-BASED APPROACH FOR IDENTIFYING BAD MOUTHING 65
6.1. Introduction 65

6.2. Solution to identify bad mouthing 66
6.2.1. Step 1: Calculate the suspicion score of reviewers 66
6.2.2. Step 2: Calculate the suspicion score of a review 68
6.2.3. Step 3: Calculate the suspicion score of product k 68
6.2.4. Step 4: Apply naïve Bayes to identify bad mouthing 69
6.3. Conclusion 71
CHAPTER 7: NETWORK-BASED APPROACH FOR IDENTIFYING SPAMMER GROUPS 73
7.1. Introduction 73
7.2. Solution to identify spammer groups 73
7.2.1. A K-means approach to identify spammer groups 74
7.2.2. A fuzzy approach to identify spammer groups 80
7.3. Conclusion 87
7.4. References 88
CHAPTER 8: GRAPH-BASED METHOD FOR DETECTING CLIQUES 89
8.1. Introduction 89
8.2. Solution to identify cliques 89
8.2.1. Step 1: Check whether a review is ballot stuffing, then calculate the total number of stars in all the reviews 89
8.2.2. Step 2: Find the shortest path distance between two nodes and calculate the influence of each node 90
8.2.3. Step 3: Define the influence region and direct region for each node 92
8.2.4. Step 4: Calculate the density of each node 92
8.2.5. Step 5: Identify and select all core nodes from node list 93
8.2.6. Step 6: Cluster the list of core nodes into separated subsets 93
8.2.7. Step 7: Identify cliques – strong core clusters 93
8.2.8. Step 8: Identify semi-strong core clusters 94
8.2.9. Step 9: Calculate providers' reputation based on detected strong core clusters (cliques) and semi-strong core clusters 95
8.3. Conclusion 95
8.4. References 95
CHAPTER 9: VALIDATION AND VERIFICATION 96

9.1. Introduction 96
9.2. Overview of Solution Implementation 96
9.2.1. Overview of solution implementation of the method to identify ballot stuffing 96
9.2.2. Overview of solution implementation of the method to identify bad mouthing 96
9.2.3. Overview of solution implementation of the method to identify spammer groups 97
9.3. Solution implementation 97
9.3.1. Solution implementation for the algorithm to identify ballot stuffing 97
9.3.2. Solution implementation for the algorithm to identify bad mouthing 111
9.3.3. Solution implementation for the algorithm to identify spammer groups 124
9.4. Conclusion 129
9.5. References 132
CHAPTER 10: RECAPITULATION AND FUTURE WORK 134
10.1. Introduction 134
10.2. Problems addressed in this thesis 134
10.3. Contribution of this thesis to the existing literature 135
10.3.1. Contribution 1 – Holistic research that identify all four types of incorrect reviews 135
10.3.2. Contribution 2 – Intelligent method to identify ballot stuffing 135
10.3.3. Contribution 3 – Intelligent method to identify bad mouthing 136
10.3.4. Contribution 4 – Intelligent methods to identify spammer groups 136
10.3.5. Contribution 5 – Intelligent method to identify cliques 136
10.4. Conclusion and future work 137
BIBLIOGRAPHY 139

LIST OF FIGURES

Figure 1: Pictorial representation of a spammer group 12
Figure 2: Pictorial representation of a clique 13
Figure 3: Example of internal factors 15
Figure 4: Example of external factors 15
Figure 5: Research methodology (Peffers et al. 2007) 44
Figure 6: The relationship between the four algorithms 47
Figure 7: A review graph (ballot stuffing) 48
Figure 8: Algorithm to identify ballot stuffing 49
Figure 9: A review graph (bad mouthing) 50
Figure 10: Algorithm to identify bad mouthing 51
Figure 11: Algorithm to identify spammer groups 53
Figure 12: Algorithm to identify cliques 55
Figure 13: Clique with numbers of reviews 90
Figure 14: Accuracy of BRS, TRAVOS, iCLUB and our method with different percentages of ballot stuffing 100
Figure 15: Precision of BRS, TRAVOS, iCLUB and our method with different percentages of ballot stuffing 101
Figure 16: Recall of BRS, TRAVOS, iCLUB and our method with different percentages of ballot stuffing 101
Figure 17: FPR and FNR of BRS with different percentages of ballot stuffing 103
Figure 18: FPR and FNR of TRAVOS with different percentages of ballot stuffing 103
Figure 19: FPR and FNR of iCLUB with different percentages of ballot stuffing 104
Figure 20: FPR and FNR of our method with different percentages of ballot stuffing 104
Figure 21: FPR and FNR of BRS, TRAVOS, iCLUB and our method when the majority are ballot stuffing (80%) 105
Figure 22: Correlation between actual reputation and admin reputation (10% ballot stuffing) 108
Figure 23: Correlation between actual reputation and admin reputation (20% ballot stuffing) 109
Figure 24: Correlation between actual reputation and admin reputation (50% ballot stuffing) 109
Figure 25: Correlation between actual reputation and admin reputation (80% ballot stuffing) 110
Figure 26: Correlation between reputation (admin) and reputation (actual) of agent Box 110
Figure 27: Accuracy of BRS, TRAVOS, iCLUB and our method with different percentages of bad mouthing 113

Figure 28: Precision of BRS, TRAVOS, iCLUB and our method with different percentages of bad mouthing 114
Figure 29: Recall of BRS, TRAVOS, iCLUB and our method with different percentages of bad mouthing 114
Figure 30: FPR and FNR of BRS with different percentages of bad mouthing 116
Figure 31: FPR and FNR of TRAVOS with different percentages of bad mouthing 116
Figure 32: FPR and FNR of iCLUB with different percentages of bad mouthing 117
Figure 33: FPR and FNR of our method with different percentages of bad mouthing 117
Figure 34: FPR and FNR of BRS, TRAVOS, iCLUB and our method when the majority is bad mouthing (80%) 118
Figure 35: Correlation between actual reputation and admin reputation (10% bad mouthing) 121
Figure 36: Correlation between actual reputation and admin reputation (20% bad mouthing) 122
Figure 37: Correlation between actual reputation and admin reputation (50% bad mouthing) 122
Figure 38: Correlation between actual reputation and admin reputation (80% bad mouthing) 123
Figure 39: Correlation between actual reputation and admin reputation of agent Box with different percentage of bad mouthing 123
Figure 40: Test dataset 125
Figure 41: Clusters of reviewers 126
Figure 42: Clusters using fuzzy K-means clustering 126

LIST OF TABLES

Table 1: Capabilities of existing approaches compared to our method 26
Table 2: Confusion matrix 98
Table 3: Confusion matrix 111
Table 4: Number of groups and number of reviewers detected using the K-means approach.....128
Table 5: Number of groups and number of reviewers detected using the fuzzy approach 128
Table 6: Number of spammer groups detected using the K-means approach and the fuzzy approach 128

ABSTRACT

With the widespread use of information technology in business, an increasing number of companies are looking for ways to reduce their overheads. The cost of IT development has been a barrier for both medium and large companies. In order to reduce cost, a new technology called cloud computing is used in many companies without private servers. When selecting which cloud provider to go with in the future, some enterprises will check certain websites for cloud reviews to see what previous cloud consumers thought about the various cloud providers. These kinds of reviews will affect their selection of a cloud provider. Therefore, the reliability of cloud reviews is very important to a cloud consumer so that they can choose a trustworthy cloud provider. In this thesis, four kinds of incorrect reviews in reputation systems are presented, namely ballot stuffing, bad mouthing, spammer groups and cliques. Previous studies on how to identify these types of incorrect reviews are also assessed in this study. Then, new methods to identify incorrect reviews, including ballot stuffing, bad mouthing, spammer groups and cliques are proposed. Finally, the solutions to identify ballot stuffing, bad mouthing and spammer groups are then validated.

THESIS SUMMARY

Reputation systems provide a method for supporting and building trust amongst different parties in online environments. The main concept of a reputation system is to let an agent rate the performance of other agents. We use the term 'agent' loosely here and it may refer to a software agent, web service or a product. A reputation score is then derived based as a mathematical function of ratings on a given agent, which will be used by other agents to determine whether or not to transact with that agent later on. Reputation systems are different from trust referral system, in which agents exchange general recommendations about other agents (Josang et al. 2003; Yolum & Singh 2003).

In reputation systems, agents rate another agent to describe their experience of a particular transaction with that agent. Therefore, the previous ratings are important as they give advice to strangers who want to interact with each other (or with a new agent) for the first time. Reputation systems are effective in providing motivation for honest behaviour and preventing dishonest behaviour amongst the agents (Buchegger & Le Boudec 2003). Finding ways to decrease or avoid unfairly high ratings, unfairly low ratings, spammer groups and cliques (or groups of agents) writing fake reviews is a fundamental issue in reputation systems, when ratings from other agents are being considered. This is because the trusting agents cannot control the reliability of the ratings when these ratings are provided by agents which are out of its control.

To make sure that a reputation system is robust, there is a need for a strong and effective mechanism to protect against unfair ratings. Furthermore, there is a need for intelligent mechanisms or methods to identify spammer groups and cliques giving fake reviews to an agent (Whitby et al. 2004).

The purpose of this research is to develop intelligent methods and algorithms to identify four types of incorrect reviews. For the scope of this research, we focus on four types of incorrect reviews as follows:

- Ballot stuffing: a seller compromises with a number of buyers to ask them to give him an unfairly high rating. In doing so, these buyers inflate the seller's reputation

which can result in increased customers and orders for products at a higher price (Dellarocas 2000).

- Bad mouthing: a seller compromises with a number of buyers to ask them to give unfairly low ratings to its enemy. In doing so, these buyers damage the reputation of the seller's enemy which can result in significantly reduced customers and orders (Dellarocas 2000).
- Spammer group: a group of reviewers who work together to write fake reviews on target products to promote or demote these products (Mukherjee et al. 2011).
- Clique: a group of agents who work together to promote their products. Agents in a clique form a review circle as one agent will receive a review from another agent and will also provide a review on a different agent.

Firstly, this thesis reviews the current literature on ways to identify these kinds of incorrect reviews on reputation systems. Secondly, the research issues, research questions and research objectives are introduced. Then, several new solutions are presented to identify ballot stuffing, bad mouthing, spammer groups and cliques. The solutions to identify ballot stuffing, bad mouthing and spammer groups are then validated, after which the thesis concludes with the research contributions and research plan.